

การจำแนกภาพพระเครื่องเบญจภาคีด้วยดีฟเลิร์นนิ่ง

Classification of Benjapakee Buddha amulets image by deep learning

นรุตม์ บุตรพลอย^{1*} และ แสงทอง บุญยิ่ง²

Narut Butploy^{1*} and Sangthong Boonying²

บทคัดย่อ

พระเครื่องเป็นสัญลักษณ์วัตถุทางความเชื่อและศาสนาที่มีความขลังและหาได้ยาก บางองค์เห็นได้เฉพาะภาพถ่าย ผู้ที่ไม่ได้ศึกษาจะไม่สามารถทราบชื่อหรือจำแนกองค์พระเครื่องได้ ปัจจุบันการรู้จำภาพด้วยเทคโนโลยีการเรียนรู้ของเครื่องจักรพัฒนาขึ้นอย่างรวดเร็ว ดีฟเลิร์นนิ่งเป็นหนึ่งในเทคโนโลยีที่ช่วยให้คอมพิวเตอร์สามารถรู้จำภาพได้อย่างแม่นยำ การพัฒนาโมเดลที่เหมาะสมกับการรู้จำภาพจะช่วยสร้างระบบอัตโนมัติสำหรับจำแนกองค์พระเครื่องเบญจภาคีได้ งานวิจัยนี้จึงเสนอการประยุกต์ใช้ดีฟเลิร์นนิ่งแบบ convolutional neural network (CNNs) สำหรับจำแนกภาพพระเครื่องเบญจภาคี วัตถุประสงค์เพื่อพัฒนาสถาปัตยกรรมดีฟเลิร์นนิ่งที่เหมาะสมกับการรู้จำองค์พระเครื่องเบญจภาคีให้ระบบสามารถจำแนกภาพได้อย่างแม่นยำให้กับผู้สนใจที่ไม่ทราบและไม่สามารถจำแนกองค์พระเครื่องแต่ละองค์ได้ การทดลองให้เครื่องคอมพิวเตอร์เรียนรู้ภาพพระเครื่องเบญจภาคีแต่ละองค์ ๑๐๐ ภาพ รวมทั้งหมด 500 ภาพ และทดลองจำแนกภาพพระเครื่องที่นำเข้ามาใหม่จำนวน 50 ภาพ จากการทดลองพบว่าสถาปัตยกรรม convolution ขนาด 3 ชั้น ให้ผลลัพธ์ที่ดีที่สุด การวัดค่าประสิทธิภาพสามารถจำแนกภาพพระสมเด็จ ภาพพระนางพญา ภาพพระรอดและภาพพระซุ้มกอถูกต้อง 80 เปอร์เซ็นต์ ส่วนภาพพระผงจำแนกถูกต้อง 70 เปอร์เซ็นต์ ซึ่งความแตกต่างระหว่างภาพองค์พระเครื่องเบญจภาคีที่เป็นเอกลักษณ์แตกต่างกันอย่างชัดเจนทำให้สามารถจำแนกภาพได้อย่างแม่นยำ แต่ถึงอย่างไรความแม่นยำจะขึ้นอยู่กับจำนวนภาพที่นำมาเรียนรู้ซึ่งโมเดลที่นำเสนอจะเหมาะสมกับภาพที่ผู้วิจัยนำมาทดสอบเท่านั้น ดังนั้นผู้ที่สนใจในการรู้จำภาพอาจจะต้องพัฒนาโมเดลที่เหมาะสมกับภาพที่นำมาทดลองต่อไป

คำสำคัญ: ดีฟเลิร์นนิ่ง สถาปัตยกรรม CNNs

Abstract

Buddha amulet is a symbol of belief and religion, which has a magical power and rare sacred object. Some Buddha amulets can only be seen by those who have not studied cannot know the name or identify the Buddha amulet. Nowadays, the image recognition using machine learning technology has developed rapidly. Deep learning is one of the technology that helps computers to recognize images precisely. The development of appropriate model for the image recognition might lead to create an automatic system for classification of Benjapakee Buddha amulets.

¹ คณะเทคโนโลยีอุตสาหกรรม มหาวิทยาลัยราชภัฏกำแพงเพชร

¹ Faculty of Industrial Technology, Kamphaeng Phet Rajabhat University

² คณะบริหารธุรกิจและเทคโนโลยีสารสนเทศ มหาวิทยาลัยเทคโนโลยีราชมงคลสุวรรณภูมิ

² Faculty of Business Administration and Information Technology, Rajamangala University of Technology Suvarnabhumi

* Corresponding author. E-mail: narut@kpru.ac.th

The research presented the applied Convolutional Neural Network system (CNN) model of deep learning to classify the Benjapakee Buddha Amulets images. The aim of the research was to develop this suitable architecture deep learning for recognizing the Benjapakee Buddha amulets accurately for the people who did not know and identify these amulets. The experiment consisted of the collection of 100 amulet images for each type, a total of 500 images, and then the test was conducted with a fitty new sample. The results showed that the 3-layer convolution architecture produced the best results. The efficiency of this model could correctly identify 80% of Phra Somdej, Phra Nang Phaya, Phra Rod and Phra Somgor, and 70% of Phra Phong images. The differences between the images of the Benjapakee Buddha amulet of the Buddha image have distinct characteristics, making the ability to accurately identify the images. However, the accuracy of the test depended on the number of tested sample images. The proposed model may only be suitable for images tested, so those interested in image recognition may have to further develop a model suitable for the images.

Keywords: deep learning, CNNs architecture

บทนำ

พระเครื่องเบญจภาคีจัดว่าเป็นสุดยอดพระเครื่องของเมืองไทยที่เป็นที่นิยมมากของวงการพระเครื่อง ได้รับความนิยมและเป็นสุดยอดปรารถนาของนักเล่นพระ และเป็นสุดยอดของพุทธคุณทาง คลาดแคล้ว คงกระพันชาตรี เมตตามหานิยม โชคลาภเงินทอง (Phinitsap, 2012) เชื่อกันว่าผู้ใดได้ครอบครองเป็นเจ้าของผู้นั้นเปี่ยมด้วยวาสนาบารมี พระเครื่องเบญจภาคี ประกอบไปด้วย พระสมเด็จวัดระฆัง จังหวัดกรุงเทพมหานคร เป็นพระประธาน พระนางพญาจังหวัดพิษณุโลก สถิตย์อยู่เบื้องขวาของสรวัยพระประธาน พระรอดจังหวัดลำพูน

สถิตย์อยู่เบื้องซ้ายของสรวัยพระประธาน พระซุ้มกอ จังหวัดกำแพงเพชร สถิตย์ถัดจากพระนางพญา และพระมงกุฎพระนคร จังหวัดสุพรรณบุรี สถิตย์ถัดจากพระรอด (Thai Buddha Image Admiration Association, 2019) เนื่องจากองค์พระเครื่องแต่ละองค์เป็นของหายาก ทำให้หาองค์พระของจริงมาศึกษายาก จึงมีการถ่ายภาพองค์พระเก็บไว้ให้เยาวชนรุ่นหลังได้ศึกษาเท่านั้น แต่เนื่องจากองค์พระแต่ละองค์แบ่งออกเป็นพิมพ์ต่างๆ ดังนั้นจำเป็นที่จะต้องอาศัยผู้เชี่ยวชาญในการดูว่าเป็นองค์พระใด ตัวอย่างภาพพระเครื่องเบญจภาคีดัง (Figure 1)

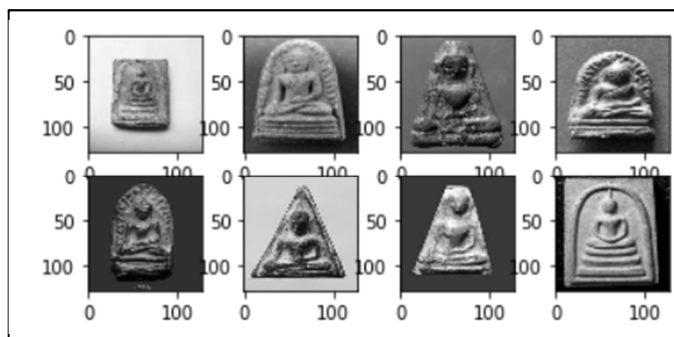


Figure 1 Example of Benjapakee Buddha amulet images.

จาก (Figure 1) ตัวอย่างภาพพระเครื่องเบญจภาคีที่นำเข้ามาเก็บไว้ในอะเรย์ของภาพเพื่อเรียกไปใช้งาน โดยแต่ละภาพจะมีป้ายบอกชื่อและหมายเลขคลาสของภาพอยู่เพื่อใช้จำแนกภาพ แต่เนื่องจากการบ่งบอกประเภทและชื่อของพระเครื่องนั้นเป็นเรื่องยากเพราะต้องอาศัยความชำนาญ ความสนใจ และประสบการณ์ของมนุษย์เป็นสำคัญ (Korbuakaew, 2007)

ปัจจุบันเทคโนโลยีด้านการเรียนรู้ของเครื่องถูกพัฒนาไปจนทำให้คอมพิวเตอร์สามารถจำแนกภาพถ่ายได้ด้วยการเรียนรู้จากภาพเรียกว่าการทำดีฟเลิร์นนิง (deep learning) ซึ่งประกอบด้วยขั้นตอนการเรียนรู้ด้วยโครงข่ายประสาทเทียมหลายชั้น งานวิจัยของ (Noord, & Postma, 2016) เสนอวิธีการทำดีฟเลิร์นนิงเพื่อจำแนกภาพหลายขนาดเพื่อสนับสนุนการรับรู้คุณสมบัติสำหรับการจัดหมวดหมู่ภาพวาด ผลการวิจัยพบว่าการทำดีฟเลิร์นนิงหลายสถาปัตยกรรมแบบ ensemble เพื่อช่วยจำภาพให้ผลลัพธ์ที่ดีกว่าการทำดีฟเลิร์นนิงแบบเดียว สามารถรู้จำและจำแนกภาพได้อย่างมีประสิทธิภาพ นอกจากนี้ งานวิจัยที่นำดีฟเลิร์นนิงมาช่วยงานด้านอื่นๆ เช่นงานด้านการแพทย์ซึ่งเป็นความท้าทายโดยงานวิจัยของ (Orlando, Prokofyeva, del Fresno, & Blaschko, 2017) นำเสนอวิธีการตรวจหารอยแดงที่ตาที่อาจจะบ่งบอกถึงโรคเบาหวานโดยอาศัยการรวมดีฟเลิร์นนิงกับ domain knowledge คุณลักษณะที่ได้จากเครือข่ายประสาทเทียมผลลัพธ์ที่ได้พบว่าการรวมแหล่งข้อมูลทั้งสองมีการปรับปรุงผลการดำเนินงานได้ดีกว่าการแยกกัน งานวิจัยของ (Sivaramakrishnan, Antani, & Jaeger, 2017) นำเสนอการนำดีฟเลิร์นนิงมาช่วยจำแนกภาพทางการแพทย์ด้วยการตรวจคัดกรองและจำแนกภาพเซลล์เชื้อมาลาเรียด้วยดีฟเลิร์นนิงโดยการ

visualizing features, visualizing activations และปรับแต่งโมเดล ซึ่งผลลัพธ์สามารถจำแนกภาพได้ถูกต้อง 98.60 เปอร์เซนต์

จากที่กล่าวมาจะเห็นว่าดีฟเลิร์นนิงสามารถนำมาช่วยจำแนกภาพได้อย่างมีประสิทธิภาพ ดังนั้นงานวิจัยนี้ผู้วิจัยจึงต้องการประยุกต์การทำดีฟเลิร์นนิงมาช่วยในการจำแนกภาพพระเครื่องเบญจภาคีโดยพระเครื่องแต่ละองค์ใช้ภาพสำหรับการเรียนรู้ 100 ภาพรวมทั้งสิ้น 500 ภาพ และนำภาพพระเครื่องที่ระบบไม่รู้จักมาทดสอบหาประสิทธิภาพวัตถุประสงค์เพื่อพัฒนาสถาปัตยกรรมดีฟเลิร์นนิงที่เหมาะสมกับการรู้จำภาพถ่ายพระเครื่องเบญจภาคีประโยชน์ของการทดลองเพื่อได้ระบบรู้จำภาพถ่ายพระเครื่องที่สามารถจำแนกภาพองค์พระเครื่องเบญจภาคีได้อย่างแม่นยำซึ่งมีขอบเขตของวิจัยคือจำแนกเฉพาะภาพถ่ายพระเครื่องเบญจภาคีโดยภาพที่นำมาทดลองเป็นเพียงภาพถ่ายจำลององค์พระเครื่องซึ่งได้จากนักสะสมพระเครื่องในจังหวัดกำแพงเพชรและการพัฒนาสถาปัตยกรรมดีฟเลิร์นนิงนี้จะต้องสามารถเรียนรู้ภาพได้อย่างถูกต้องไม่น้อยกว่า 90 เปอร์เซนต์ขึ้นไป

วิธีการศึกษา

งานวิจัยนี้ใช้หลักการการประมวลผลภาพด้วยดีฟเลิร์นนิงซึ่งมีหลักการโครงข่ายประสาทเทียมแบบหลายชั้น (multi-layer perceptron: MLP) ซึ่งมีแนวคิดคือให้คอมพิวเตอร์เรียนรู้หลายขั้นตอนในลักษณะขนาน ยิ่งเกิดการเรียนรู้มากและลึกขึ้นคอมพิวเตอร์จะสามารถทำนายข้อมูลที่เข้ามาได้ถูกต้อง การประมวลผลด้วยดีฟเลิร์นนิงนี้จะต้องมีทรัพยากรรองรับได้แก่หน่วยประมวลผลกลาง (CPU)

ที่มีแบบหลายแกนและการ์ดจอ (GPU) ที่มีแกนประมวลผลแบบหลายแกน เนื่องจากการเรียนรู้จะส่งคำสั่งไปประมวลผลในลักษณะขนาน อัลกอริทึมแบบดีฟเลิร์นนิ่งเป็นเซตของอัลกอริทึมการเรียนรู้ของเครื่องซึ่งมีภายในการทำงานของโครงข่ายประสาทเทียมหลายเลเยอร์เพื่อแก้ปัญหาปัญญาประดิษฐ์แบบดั้งเดิม (Remez, Litany, Giryas, & Bronstein, 2017) โดยโมเดลที่เหมาะสมสำหรับการจำแนกภาพได้แก่ CNNs ซึ่งเป็นหนึ่งในวิธีการเรียนรู้ที่มีประสิทธิภาพและเป็นที่นิยมใช้ในการรู้จำและจำแนกภาพให้กับคอมพิวเตอร์ (LeCun, Bengio, & Hinton, 2015) ซึ่งสถาปัตยกรรม CNNs เป็นโครงข่ายประสาทเทียมที่ประกอบด้วยชั้น convolution ซึ่งเป็นการสกัดลักษณะสำคัญต่างภาพออกมา การสกัดลักษณะของภาพออกมานี้จะถูกปรับไปพร้อมกับส่วนจำแนก ดังนั้นหลังจากการเรียนรู้ภาพแล้วจะได้คุณลักษณะที่เหมาะสมกับงาน เซลล์ประสาทในชั้นนี้จะถูกจัดเป็นคุณลักษณะ (feature map) โดยจะมีฟิลต์ซึ่งเชื่อมต่อกับเซลล์ประสาทในชั้นก่อนหน้าเซลล์ประสาททั้งหมดภายในคุณลักษณะมีน้ำหนักที่เท่ากัน ในบางครั้งผู้ออกแบบสถาปัตยกรรม convolutional จะนำชั้น pooling มาแทรกระหว่างชั้น convolution ซึ่งหน้าที่ของ pooling คือการลดขนาดเชิงพื้นที่เพื่อลดจำนวนของพารามิเตอร์และการคำนวณในเครือข่าย

ซึ่งงานวิจัยนี้ผู้วิจัยจะทดลองสร้างสถาปัตยกรรมดีฟเลิร์นนิ่งแบบ CNNs ที่เหมาะสมกับการเรียนรู้และจำแนกภาพพระเครื่องเบญจภาคีโดยมีรายละเอียดการทดลองดังนี้

ชุดข้อมูล

การที่จะให้ดีฟเลิร์นนิ่งสามารถจำแนกภาพได้อย่างแม่นยำจำเป็นต้องเรียนรู้จากภาพสำหรับฝึกฝนระบบจำนวนมาก งานวิจัยนี้ได้นำภาพพระเครื่องเบญจภาคีจากนักสะสมพระเครื่องในจังหวัดกำแพงเพชร แต่เนื่องจากพระเครื่องเบญจภาคีองค์จริงนั้นเป็นวัตถุมงคลที่ล้ำค่าไม่สามารถถ่ายภาพองค์จริงได้ทำให้ภาพที่นำมาทดลองจะเป็นองค์จำลอง ดังนั้นงานวิจัยนี้จะเป็นการสร้างสถาปัตยกรรมการเรียนรู้เพื่อจำแนกภาพพระเครื่องแต่ละองค์ซึ่งยังต้องพัฒนาต่อในการสร้างสถาปัตยกรรมดีฟเลิร์นนิ่งสำหรับจำแนกองค์พระเครื่องจริงกับองค์พระเครื่องจำลอง การทดลองจะแบ่งภาพองค์พระเครื่องเป็น 2 ชุด ได้แก่ ชุดฝึกฝนระบบจำนวน 500 ภาพ ซึ่งมีภาพต้นฉบับแต่ละองค์จำนวน 20 ภาพ นำมาทำ image augmentation ด้วยการหมุนภาพ ทำ horizontal และ vertical flip จะได้ภาพที่แตกต่างกันอีก 5 ภาพ ดังนั้นภาพพระเครื่องแต่ละองค์จะมีภาพทั้งหมด 100 ภาพ และเตรียมชุดสำหรับทดสอบระบบ ซึ่งจะคละภาพองค์พระเครื่องทั้ง 5 องค์ ที่ไม่ได้ใช้สอนระบบองค์ละ 10 ภาพ รวม 50 ภาพ

สถาปัตยกรรม CNNs

การมองภาพของ CNNs จะมองสิ่งที่เข้ามาเป็น 3 มิติ ซึ่งมี volume เป็นความกว้าง ยาว ลึก การออกแบบสถาปัตยกรรมดีฟเลิร์นนิ่งแบบ CNNs จะขึ้นอยู่กับความเหมาะสมของงานและผู้ใช้งานซึ่งโดยปกติแล้วสถาปัตยกรรมของ CNNs จะประกอบไปด้วยเลเยอร์หลัก จำนวน 3 เลเยอร์ ดัง (Figure 2)

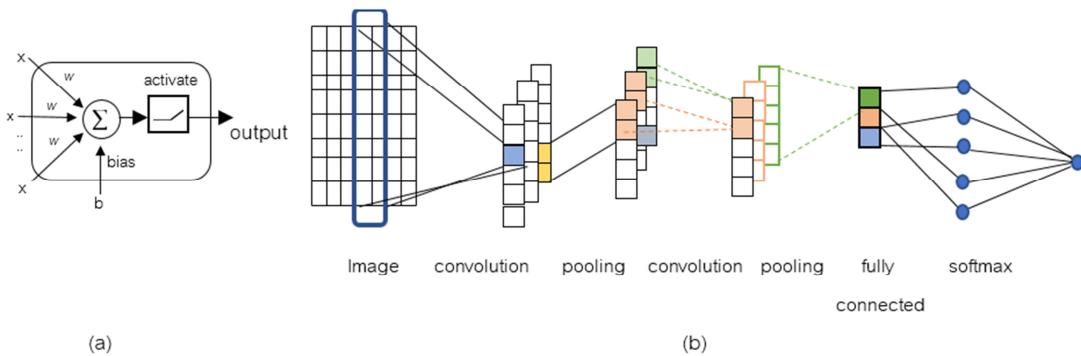


Figure 2 The CNNs architecture.

จาก (Figure 2b) คือสถาปัตยกรรมการประมวลผลภาพแบบ CNNs ซึ่งประกอบไปด้วยชั้น convolutional ที่เป็นการนำตัวกรอง (kernel) ขนาด $n \times n$ มาประมวลผลภาพเพื่อนำคุณลักษณะที่สำคัญของภาพออกมา ชั้น pooling เป็นการลดขนาดภาพให้เล็กลงเนื่องจากคุณลักษณะที่สำคัญที่ถูกนำขึ้นมานั้น มีด้วยกันหลายมิติซึ่งต้องใช้ทรัพยากรของเครื่องคอมพิวเตอร์ในการ คำนวณค่อนข้างมาก ดังนั้นเพื่อเป็นการลดการคำนวณจำเป็นต้องลดขนาดภาพซึ่งการทำ pooling จะหาค่าที่ให้สารสนเทศสูงสุดของตำแหน่งที่ใกล้เคียงกัน ดังนั้นจะลดตำแหน่งที่ไม่ถูกนำมาคำนวณทำให้การคำนวณได้รวดเร็วขึ้น ขั้นสุดท้ายชั้น fully connected คือการรวมทุกคุณลักษณะที่ได้มารวมกันแล้วทำการจำแนกภาพ นอกจากนี้ในแต่ละครั้งของการสร้างสถาปัตยกรรมการเรียนรู้แบบ CNNs จะต้องคำนึงถึงการอัปเดตค่าถ่วงน้ำหนักด้วยฟังก์ชันการกระตุ้นแบบ relu เพื่อแก้ปัญหาการสั่นของระบบที่อาจเกิด gradient มีขนาดเล็กลงเรื่อยๆ ทำให้ค่าถ่วงน้ำหนักไม่ถูกอัปเดต และการแก้ปัญหาการจำคำตอบด้วยการ dropout บางโหนดซึ่งมีรายละเอียดดังนี้

ฟังก์ชันกระตุ้น การทำงานภายในของดีฟเลิร์นนิง ซึ่งเป็นโครงข่ายประสาทเทียมที่เชื่อมต่อกันระหว่างโหนด แต่ละการเชื่อมต่อนี้จะมีค่าถ่วงน้ำหนัก (weight, w) กำกับอยู่ที่เส้นเมื่อโครงข่ายเริ่มทำงานข้อมูลนำเข้า (x) จะส่งผ่านโหนดนำเข้าดัง (Figure 2b) แล้วคำนวณร่วมกับค่าถ่วงน้ำหนักและค่าไบแอส (bias, b) เมื่อคำนวณข้อมูลนำเข้าทั้งหมดแล้วก่อนส่งออกไปยังโหนดถัดไปจะต้องผ่านฟังก์ชันกระตุ้น โดยฟังก์ชันกระตุ้นที่นิยมใช้ในดีฟเลิร์นนิงแบบ CNNs ได้แก่ ฟังก์ชัน rectified linear unit : relu ซึ่งถ้าค่าที่ส่งมามากกว่าศูนย์ จะได้ผลลัพธ์จะเป็นศูนย์แต่หากค่าที่ส่งมามากกว่าหรือเท่ากับศูนย์ผลลัพธ์ที่ได้จะมากกว่าศูนย์ เมื่อเป็นเช่นนั้นถ้าค่านำเข้าเป็นบวกจะส่งผลให้ผลลัพธ์มากกว่าศูนย์ตลอดทำให้ gradient ที่เป็นค่าส่งกลับสำหรับอัปเดตโครงข่ายไม่หายทำให้โมเดลที่สร้างเรียนรู้ได้เป็นการแก้ปัญหา vanishing gradient (Krizhevsky, Sutskever, & Hinton, 2012) ดังนั้นยังสถาปัตยกรรมที่สร้างขึ้นลึกมากเท่าไรยิ่งส่งผลต่อ gradient มากขึ้นชั้นที่อยู่ลึกจะไม่ได้รับค่าการอัปเดตจากผลการเรียนรู้ดังนั้นเพื่อแก้ปัญหาจะต้องเพิ่มฟังก์ชันการกระตุ้นแบบ relu

เข้าไปในสถาปัตยกรรมที่สร้างขึ้น ฟังก์ชันกระตุ้นแบบ sigmoid ฟังก์ชันนี้เป็นฟังก์ชันที่ไม่ใช่เชิงเส้นค่าผลลัพธ์ที่ได้จากการผ่านค่าอยู่ระหว่าง 0 ถึง 1 เหมือน relu แต่การส่งผ่านค่าจะเกิดปัญหากรณีที่ค่านำเข้า น้อยกว่า -5 หรือมากกว่า 5 ความชันจะเข้าใกล้ 0 จนทำให้ gradient หายส่งผลไม่สามารถอัปเดตโครงข่ายได้และอาจเกิดปัญหา vanishing gradient

แต่เพื่อให้ได้ข้อมูลที่ถูกต้องในงานวิจัยนี้ผู้วิจัยจะทดลองใช้ฟังก์ชันกระตุ้นทั้งสองแบบในการสอนระบบเพื่อหาฟังก์ชันกระตุ้นที่เหมาะสมกับการเรียนรู้ภาพพระเครื่อง

dropout การทำงานของโครงข่ายประสาทเทียมเมื่อเรียนรู้ไปซักระยะจะทำให้โครงข่ายฉลาดและจำคำตอบไว้ด้วยการวิ่งไปหาโหนดที่เคยได้คำตอบที่

ถูกต้องปัญหานี้เรียกว่า overfitting ซึ่งผลเสียจะร้ายแรงเนื่องจากโครงข่ายจะจำเฉพาะเส้นทางที่เคยไปแล้วให้ผลลัพธ์ที่ดี แต่เมื่อเจอปัญหาใหม่หรือภาพใหม่โครงข่ายจำว่าต้องไปเส้นทางเดิมนั้นผลการทำนายจะผิดพลาดได้ เพื่อป้องกันปัญหาดังกล่าวจึงจำเป็นต้องตัดโหนดที่เคยเรียนรู้ด้วย dropout ซึ่งวัตถุประสงค์หลักคือการทำลายโหนดที่เคยเรียนรู้แล้วให้โครงข่ายพยายามหาเส้นทางใหม่แล้วเรียนรู้และสร้างโมเดลใหม่เสมอ (Srivastava, Hinton, Krizhevsky, Sutskever, & Salakhutdinov, 2014)

เมื่อทราบถึงรูปแบบและฟังก์ชันที่จำเป็นสำหรับการสร้างสถาปัตยกรรมการเรียนรู้แบบ CNNs แล้วใน (Figure 3) จะนำเสนอตัวอย่างการเรียนรู้ของสถาปัตยกรรม CNNs ดีฟเลิร์นนิ่ง

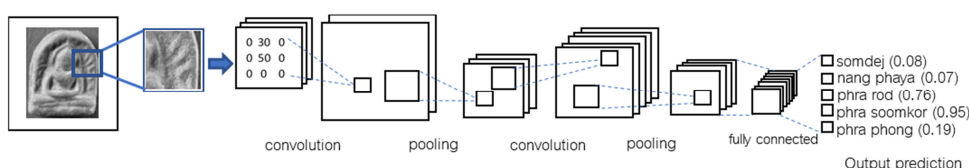


Figure 3 The example shows how to use a trained CNNs deep learning.

จาก (Figure 3) นำเสนอตัวอย่างการเรียนรู้ของ CNNs ดีฟเลิร์นนิ่ง โดยเริ่มจากระบบจะนำตัวกรองขนาด $n \times n$ มาประมวลกับภาพซึ่งตัวกรองจะมีหลายรูปแบบ เช่น identity, blur, sharpen (Figure 4) เพื่อให้ network เรียนรู้ภาพในมิติที่แตกต่างกันจากนั้นจะสร้างคุณลักษณะ

ที่ได้จากการคำนวณระหว่างภาพนำเข้ากับตัวกรอง เมื่อสร้างคุณลักษณะจะลดขนาดการคำนวณด้วยการทำ pooling ซึ่งงานวิจัยนี้ทดลองสร้างชั้น convolution และ pooling แบบนี้หลายชั้นเพื่อหาสถาปัตยกรรมที่เหมาะสมที่สุดสำหรับจำแนกภาพพระเครื่องเบญจภาคี

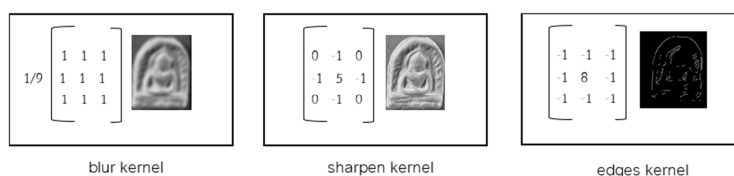


Figure 4 The example of kernel operation.

จาก (Figure 4) นำเสนอตัวกรองประเภทต่างๆ ที่เกิดขึ้นในการทำงานของ CNNs ดีฟเลิร์นนิ่งเพื่อสร้างมิติภาพการเรียนรู้ทำให้โครงข่ายจดจำภาพได้ในหลากหลายมิติ เมื่อภาพใหม่ที่เข้าจะสามารถจำแนกภาพได้มากขึ้น ในส่วนการพัฒนาโปรแกรมและทรัพยากร

ด้านฮาร์ดแวร์ผู้วิจัยใช้เครื่องคอมพิวเตอร์ที่มีหน่วยประมวลผลภาพรุ่น GTX1050 หน่วยประมวลผลกลาง intel core i7 หน่วยความจำขนาด 4 กิกะไบต์ ซอฟต์แวร์ในการพัฒนาโปรแกรมภาษา python และไลบรารี keras และออกแบบการทดลอง ดัง (Figure 5)

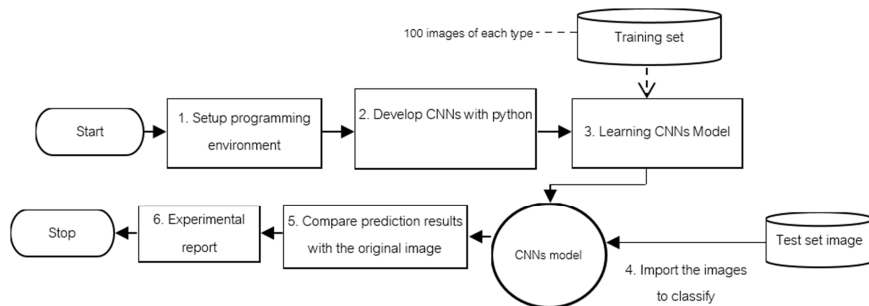


Figure 5 The experimental overview.

จาก (Figure 5) แสดงภาพรวมการทดลองโดยเริ่มจากการติดตั้งสภาพแวดล้อมของการเขียนโปรแกรมเพื่อให้ตัวแปลภาษาสามารถเรียกใช้หน่วยประมวลผลกราฟิกสำหรับประมวลผลภาพ การทดลองจะนำภาพจากคลังภาพจำนวน 500 ภาพ แบ่งออกเป็น 5 คลาส แต่ละคลาสจะประกอบด้วยภาพพระเครื่องเบญจภาคี

องค์ละ 100 ภาพมาเข้าโมเดลสถาปัตยกรรมที่สร้างไว้เมื่อเรียนรู้ภาพแล้วจะทดลองนำภาพที่ต้องการจำแนกภาพซึ่งเป็นภาพใหม่ที่ไม่อยู่ใน 500 ภาพข้างต้น เข้าโมเดลแล้วให้โปรแกรมจำแนกออกมาว่าเป็นภาพพระเครื่องเบญจภาคีองค์ใด ในส่วนสถาปัตยกรรมที่สร้างขึ้นนำเสนอไว้ดัง (Figure 6)

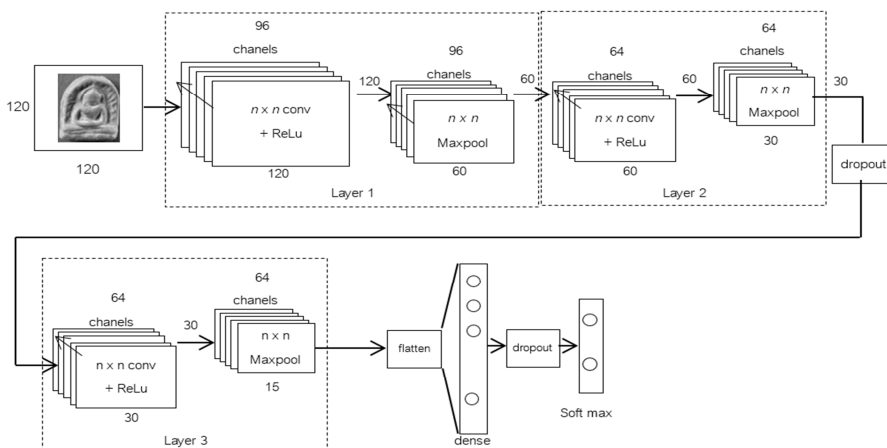


Figure 6 The proposed architecture.

จาก (Figure 6) เป็นสถาปัตยกรรม CNNs ที่นำเสนอในงานวิจัย โดยชั้นเรียนรู้แรกจะประกอบไปด้วยการตรวจจับคุณลักษณะภาพ ทดลองเปลี่ยนขนาด kernel size $n \times n$ ตั้งแต่ 2×2 , 3×3 และ 4×4 แต่ละ size ใช้ฟังก์ชันกระตุ้น relu และ sigmoid ทดลองทั้งหมด 5 ชั้น งานวิจัยของ Lin, Chen, & Yan (2014) ได้นำเสนอสถาปัตยกรรมแบบ CNNs โดยใช้วิธี multilayer perceptron และ global average pooling layer รวมกัน งานวิจัยของ Springenberg, Dosovitskiy, Brox, & Riedmiller (2015) เป็นการนำเสนอแบบจำลองที่มีลักษณะเป็นรูปแบบทั่วไปแต่ที่ต่างไปคือชั้นสุดท้ายเชื่อมต่อกับ global average pooling layer งานวิจัยของ Krizhevsky, Sutskever, & Hinton (2012) ได้ทดสอบทำ fully connected 3 ครั้ง ผู้วิจัยจึงใช้แนวทางการกำหนดชั้นการเรียนรู้จากที่กล่าวมา ในส่วนการวัดประสิทธิภาพของสถาปัตยกรรม ผู้วิจัยใช้ confusion matrix ขนาด n แถว และ n คอลัมน์ และหาค่าความแม่นยำ (precision) ค่าระลึก (recall) และค่าประสิทธิภาพโดยรวมของระบบ (F-measure) โดยค่าทั้ง 3 สามารถหาได้ดังนี้ (Phawattanakul, 2012) $\text{precision} = TP / (TP + FP)$, $\text{recall} = TP /$

$(TP + FN)$ และ $f \text{ measure} = 2 \times ((\text{precision} \times \text{recall}) / (\text{precision} + \text{recall}))$

ผลการศึกษา

เพื่อหาประสิทธิภาพของสถาปัตยกรรมผู้วิจัยได้แยกการทดลองไว้ 2 ส่วน ส่วนแรกเป็นการทดสอบหาประสิทธิภาพของสถาปัตยกรรมที่นำเสนอด้วยการหาค่าความถูกต้องของการเรียนรู้ภาพของโปรแกรม ส่วนที่ 2 ทำนายภาพพระเครื่องเบญจภาคีที่มีภาพคละภาพกันเก็บไว้ไม่ได้เรีกทหริเดียวกันจำนวน 50 ภาพ

1. ผลการทดสอบหาประสิทธิภาพของสถาปัตยกรรมที่นำเสนอ ในการทดสอบนี้ผู้วิจัยได้ทดลองเขียนโปรแกรมโดยแบ่งงานออกเป็น 2 ส่วนย่อย ได้แก่ ส่วนที่ 1 หาค่าความถูกต้องของการเรียนรู้เพื่อหาสถาปัตยกรรมที่ให้ผลลัพธ์การเรียนรู้ที่ดีที่สุด โดยกำหนดให้สถาปัตยกรรมมีชั้นของฟังก์ชันกระตุ้นแบบ relu ร่วมกับชั้น dropout, ชั้นของฟังก์ชันกระตุ้นแบบ relu ร่วมกับชั้นที่ไม่มี dropout, ชั้นของฟังก์ชันกระตุ้นแบบ sigmoid ร่วมกับชั้น dropout, ชั้นของฟังก์ชันกระตุ้นแบบ sigmoid ร่วมกับชั้นที่ไม่มี dropout เพื่อหาค่าความถูกต้องของการเรียนรู้ที่ดีที่สุด

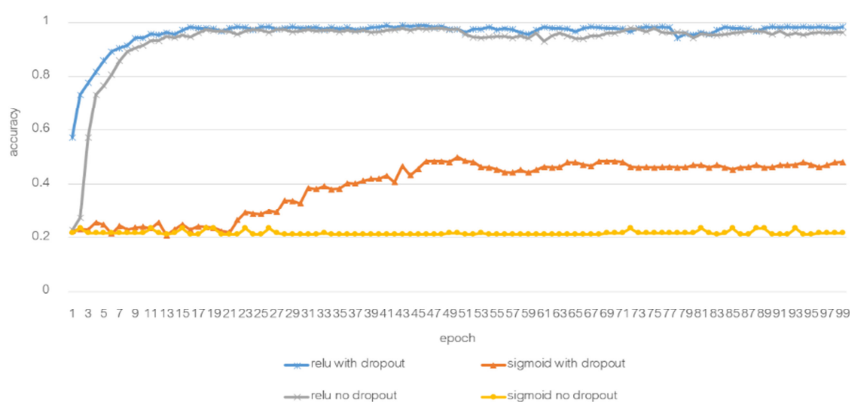


Figure 7 The accuracy of relu with dropout, relu without dropout, sigmoid with dropout, and sigmoid without dropout.

จาก (Figure 7) นำเสนอการเปรียบเทียบผลการเรียนรู้ทั้ง 4 แบบ ผลลัพธ์ที่ได้พบว่าตัวแบบที่มีชั้นของฟังก์ชันกระตุ้นแบบ relu ร่วมกับชั้น dropout และชั้นของฟังก์ชันกระตุ้นแบบ relu ร่วมกับชั้นที่ไม่มี dropout ให้ผลลัพธ์การเรียนรู้ที่มากกว่าฟังก์ชันกระตุ้นแบบ sigmoid ดังนั้นในการทดลองนี้จึงเลือกที่จะสร้างสถาปัตยกรรม CNNs ด้วยชั้นของฟังก์ชันกระตุ้นแบบ relu ร่วมกับชั้น dropout เนื่องจากให้ผลลัพธ์การเรียนรู้

ที่ดีและป้องกันการเกิด overfitting ที่โมเดลจะจำเฉพาะโหนดที่ถูกต้องตลอดเวลาทำให้เมื่อมีภาพเข้ามาใหม่อาจเกิดการเรียนรู้ที่ผิดพลาดได้

เมื่อได้สถาปัตยกรรมที่ต้องการแล้วผู้วิจัยได้นำสถาปัตยกรรมที่ได้มาทดลองต่อในส่วนย่อยที่ 2 เพื่อหาจำนวนชั้นและรอบการเรียนรู้ของสถาปัตยกรรมที่ให้ผลการเรียนรู้ที่เหมาะสมที่สุดสำหรับเรียนรู้ภาพพระเครื่องเบญจภาคี

Table 1 The accuracy of images learning from proposed architecture.

	100 epoch			150 epoch			200 epoch			250 epoch			300 epoch		
	2×2	3×3	4×4	2×2	3×3	4×4	2×2	3×3	4×4	2×2	3×3	4×4	2×2	3×3	4×4
1 layer	87.61	88.29	86.47	86.47	87.88	85.29	88.23	87.05	85.29	85.29	88.47	84.71	87.05	89.23	85.35
2 layer	88.61	89.41	88.82	88.82	88.23	88.24	86.47	88.17	88.82	88.23	89.41	88.82	88.82	89.41	87.05
3 layer	86.52	91.76	88.64	84.70	90.70	83.53	83.52	91.58	84.70	85.88	90.70	85.29	87.64	90.70	88.26
4 layer	83.47	90.41	87.05	85.88	90.41	88.24	84.64	89.23	85.29	85.88	89.23	88.23	85.88	89.23	88.23
5 layer	86.47	88.47	85.88	82.94	85.29	87.05	87.05	84.70	81.17	81.76	83.53	85.88	87.05	87.05	85.88

จาก (Table 1) นำเสนอค่าความถูกต้องการเรียนรู้ภาพพระเครื่องเบญจภาคีโดยกำหนดขนาดชั้นและรอบการเรียนรู้ขนาดต่างๆ ผลการทดลองพบว่าที่จำนวนขนาดชั้นเรียนรู้ 3 ชั้น kernel size ขนาด 3×3 รอบการเรียนรู้ทุกขนาดและขนาดชั้นการเรียนรู้ 4 ชั้นรอบการเรียนรู้ที่ 100 และ 150 รอบ kernel size ขนาด 3×3 ให้ผลลัพธ์การเรียนรู้มากเกิน 90 เปอร์เซนต์ และสิ่งที่สังเกตพบคือเมื่อเพิ่มชั้นการเรียนรู้เป็น 5 ชั้นขึ้นไป ผลการเรียนรู้ในทุกขนาด kernel size ให้ผลลัพธ์ที่น้อยลง

2. ผลการจำแนกภาพพระเครื่องเบญจภาคีแบบรวมด้วยสถาปัตยกรรมที่นำเสนอ จากผลการทดลองที่ผ่านมาพบว่าที่ชั้นการเรียนรู้ขนาด 3 ชั้นให้ผลการทดสอบที่ดีที่สุด ดังนั้นผู้วิจัยจึงทดสอบประสิทธิภาพของสถาปัตยกรรมต่อด้วยการทดลองให้จำแนกภาพที่มีภาพพระเครื่องเบญจภาคีแต่ละองค์แบบคละไม่เรียงลำดับภาพผลการทดลองเป็นไปดัง (Table 2)

Table 2 The results of classification Benjapakee buddha amulet images.

	A	B	C	D	E	recall
A	8	0	1	1	0	80%
B	0	8	0	0	2	80%
C	1	0	8	1	0	80%
D	0	0	2	8	0	80%
E	0	2	1	0	7	70%
precision	88.88%	80%	66.66%	80%	77.77%	
f- measure	84.21%	80%	72.73%	80%	73.68%	

เมื่อ A คือ ภาพพระสมเด็จ B คือ ภาพพระนางพญา C คือ ภาพพระรอด D คือ ภาพพระซุ้มกอ และ E คือ ภาพพระผง

จาก (Table 3) แสดงผลการจำแนกภาพพระเครื่องจำนวน 50 ภาพซึ่งผลการทดลองการจำแนกภาพพบว่าภาพสมเด็จวัดระฆัง พระนางพญา พระรอด และพระซุ้มกอ ทำนายถูกต้องจริง 8 ภาพ จาก 10 ภาพ ส่วนพระผงทำนายถูกเพียง 7 ภาพ คิดค่า f-measure ของภาพสมเด็จวัดระฆังได้เท่ากับ 84.21 เปอร์เซ็นต์ ภาพพระนางพญาเท่ากับ 80 เปอร์เซ็นต์ ภาพพระรอดเท่ากับ 72.73 เปอร์เซ็นต์ ภาพพระซุ้มกอเท่ากับ 80 เปอร์เซ็นต์ ภาพพระผงเท่ากับ 73.68 เปอร์เซ็นต์ และค่าเฉลี่ยอยู่ที่ 78.12 เปอร์เซ็นต์

อภิปรายผล

จากผลการทดลอง (Figure 7) จะเห็นว่าถ้าใช้ฟังก์ชันกระตุ้นแบบ relu และทำลายบางโหนดด้วย dropout จะให้การเรียนรู้ที่ถูกต้องกว่าฟังก์ชันกระตุ้นแบบ sigmoid งานวิจัยจึงเลือกใช้ฟังก์ชันกระตุ้นแบบ relu และการทำลายโหนดด้วย dropout ในงานวิจัยและนำเสนอสถาปัตยกรรมดัง (Figure 6) ใช้ kernel size ขนาด 2×2, 3×3 และ 4×4 ทดลองสร้างชั้นการเรียนรู้จำนวน 5 ชั้น ผลลัพธ์ที่ได้พบว่า kernel size ขนาด 3×3 ที่ 3 ชั้นให้ผลลัพธ์ที่ดีที่สุดและสังเกตพบว่าจำนวนชั้นเรียนรู้มีความสัมพันธ์กับการกำหนด kernel size ดัง (Table 2) เมื่อมีชั้นเรียนรู้จำนวนเพิ่มขึ้นขนาดผลลัพธ์ขึ้น แต่เป็นที่น่าสังเกตว่าประสิทธิภาพการทำงานจะเพิ่มขึ้นในระดับหนึ่งเมื่อชั้นเรียนรู้มากขึ้นไปประสิทธิภาพจะลดลง เนื่องจากสถาปัตยกรรม CNNs ขึ้นอยู่กับจำนวนข้อมูลที่นำมาสอนระบบการ

เพิ่มขึ้นมากเกินไปจะไม่มีผลใดๆ การเรียนรู้ภาพนั้นมีส่วนสำคัญคือ เส้น ส่วนหัว ส่วนโค้ง เม็ดสี การเรียนรู้สามารถคำนวณได้ตั้งแต่ขั้นต้นๆ การเพิ่มจำนวนชั้นการเรียนรู้มากจะใช้ในกรณีชุดข้อมูลที่มีขนาดใหญ่ที่มีปัจจัยอื่นที่ช่วยในการแยกแยะภาพทำให้ต้องเพิ่มจำนวนชั้นการเรียนรู้ที่มากขึ้น ในส่วนของชั้นการเรียนรู้ที่ 3 ให้ผลลัพธ์ที่ดีที่สุดอาจเนื่องมาจากการตั้งค่าการรับรู้ภาพโดยทั่วไปเลเยอร์แรกอาจเรียนรู้องค์ประกอบที่เรียบง่ายของภาพ ขึ้นถัดไปอาจเรียนรู้คุณลักษณะที่ซับซ้อนมากขึ้นซึ่งเป็นชุดของขอบ เช่น รูปร่าง และเมื่อมาถึงชั้นที่สามารถเรียนรู้การผสมผสานรูปทรงได้ตรงกับภาพต้นฉบับ ในการทดลองนี้จะใช้เครื่องมือที่ต่างจากงานวิจัยของ Korbuakaew (2007) และ Fugthong, & Meesad, (2013) ตรงที่งานวิจัยนี้ใช้ดีฟเลิร์นนิ่งซึ่งไม่เพียงตรวจสอบเฉพาะลักษณะเด่นแต่จะตรวจสอบสีในแต่ละพิกเซลดังนั้นไม่จำเป็นต้องแปลงภาพให้อยู่ในลักษณะ gray scale เพื่อสร้างขอบ แต่ดีฟเลิร์นนิ่งมีตัวกรองของตัวเองที่สุ่มขึ้นมาแล้วเรียนรู้ไปเรื่อยๆ และแปลงทุกอย่างในภาพให้เป็นสารสนเทศเก็บไว้สำหรับเรียนรู้ภาพให้มีประสิทธิภาพการทำนายสูงกว่า

ในส่วนของประสิทธิภาพของการจำแนกภาพดัง (Table 2) พบว่าส่วนใหญ่ที่ทำนายผิดพลาดเนื่องจากภาพที่นำมาทดสอบบางภาพมีลักษณะเด่นใกล้เคียงกัน เช่น การจำแนกภาพพระผงแต่ละระบบทำนายเป็นภาพพระนางพญาสองภาพเนื่องมาจากลักษณะเป็นรูปทรงสามเหลี่ยมคล้ายกัน ทำนายเป็นพระรอด 1 ภาพ อาจเนื่องมาจากภาพที่นำมาทำนายมีกรอบที่แตกหักบิ่นคล้ายคลึงกัน ซึ่งจากประสิทธิภาพการทำงานจะพบว่าภาพต้นฉบับมีความสำคัญมาก

ทั้งกรอบ ลักษณะเด่นต่างๆ ของพระเครื่องแต่ละองค์ ดังซึ่งในงานครั้งถัดไปผู้วิจัยมีความประสงค์จะศึกษา และพัฒนาการจำแนกรายละเอียดของแต่ละพิมพ์ ของแต่ละองค์เพื่อสร้างโมเดลการเรียนรู้ที่ฉลาดในการจำแนกพิมพ์พระเครื่องต่อไป

สรุป

การทดลองนี้ผู้วิจัยนำเสนอสถาปัตยกรรม CNNs ที่เหมาะสมกับการจำแนกภาพพระเครื่อง โดยแบ่งการทดลองออกเป็น 2 ส่วน ได้แก่ การทดลองจำแนกภาพพระเครื่องเบญจภาคีแต่ละองค์ในแต่ละได้เรีกทอรี่และจำแนกภาพพระเครื่องเบญจภาคีที่ปนกันในแต่ละได้เรีกทอรี่เดียวกัน ทั้งสองการทดลองให้ผลลัพธ์ไปในทิศทางเดียวกันคือจำนวนชั้นการเรียนรู้ที่ให้ผลลัพธ์ดีที่สุดคือ 3 ชั้น kernel size ขนาด 3×3 มีข้อสังเกตคือลักษณะพิเศษขององค์พระเครื่องที่เข้ามาเรียนรู้จะช่วยให้สามารถจำแนกได้แม่นยำมากขึ้น โดยงานวิจัยของ Korbuakaew (2007) พบข้อสังเกตเดียวกันว่าระบบจะทำการจำแนกได้อย่างมีประสิทธิภาพมากขึ้นหากสามารถจัดเก็บลักษณะพิเศษเฉพาะของภาพองค์พระในแม่แบบได้มาก นอกจากนี้ผู้วิจัยยังมีแนวคิดในการพัฒนาต่อยอดการจำแนกองค์พระเครื่องในแต่ละพิมพ์และสถานที่จัดสร้างเพื่อสร้างระบบการตรวจสอบวัตถุมงคลเป็นการอนุรักษ์ศิลปวัฒนธรรมของประเทศต่อไป

เอกสารอ้างอิง

- Fugthong, P., & Meesad, P. (2013). Buddha amulet information retrieval using digital images combined with feature extraction and K-nearest neighbor techniques. *Information Technology Journal*, 9(2), 34-40. (in Thai)
- Korbuakaew, C. (2007). *Identification of amulets with special feature matching* (Master's thesis). Silpakorn University, Bangkok. (in Thai)
- Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). ImageNet classification with deep convolutional neural networks. In F. Pereira (Ed.), *NIPS'12 Proceedings of the 25th International Conference on Neural Information Processing Systems* (pp. 1097-1105). Nevada: Curran Associates Inc.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature International journal of science*, 521, 436-444.
- Lin, M., Chen, Q., & Yan, S. (2014). Network in network. In Y. Bengio (Ed.), *International Conference on Learning Representations 2014 (ICLR 2014)*. Banff: ICLR.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). *Dropout: a simple way to prevent neural networks from overfitting*. *J. Mach. Learn. Res.*, 15, 1929-1958.
- Noord, N., & Postma, E. (2016). *Learning scale-variant and scale-invariant features for deep image classification*. Retrieved 7 April 2019, from <https://arXiv:1602.01255.pdf>
- Orlando, J. I., Prokofyeva, E., del Fresno, M., & Blaschko, M. B. (2017). An ensemble deep learning based approach for red lesion detection in fundus images. *Comput Methods Programs Biomed*, 153, 115-127.
- Phawattanakul, K. (2012). *Suggestion Mining from reviewers' reviews of television programs* (Master's thesis). National Institute of Development Administration.

- Phinitap, S. (2012). *Education for the sake of creativity: Case the degeneration in the faith on the Buddha series five parties in Thai society* (Master's thesis). Srinakharinwirot University, Bangkok.
- Remez, T., Litany, O., Giryas, R., & Bronstein, A. M. (2017). *Deep convolution denoising of low-light images*. Retrieved 7 April 2019, from <https://arxiv.org/pdf/1701.01687.pdf>
- Sivaramakrishnan, R., Antani, S., & Jaeger, S. (2017). Visualizing Deep Learning Activations for Improved Malaria Cell Classification. In *Proceedings of First workshop Medical Informatics and Healthcare, Machine Learning Research vol.69* (pp. 40-47). Halifax Canada: Machine Learning Research.
- Springenberg, J. T., Dosovitskiy, A., Brox, T. & Riedmiller, M. (2015). Striving for simplicity: The all convolutional net. In Y. Bengio (Ed.), *International Conference on Learning Representations 2015 (ICLR 2015) Workshop Track*. San Diego: ICLR.
- Thai Buddha Image Admiration Association. (2019). *Genesis of the Benjapakee Buddha Amulet*. Retrieved 7 April 2019, from <https://www.samakomphra.com> (in Thai)